More on Kernel Regression

$MSE(h) = bias^2(h) + Variance(h)$ want to trade less bias for more variance until we reach minimum

$$\Rightarrow \frac{\partial MSE}{\partial h} = 0 = \frac{d(bias^2)}{dh} + \frac{d\,Variance}{dh}$$

Bias from bandwidth:

$$\begin{aligned} bias(x) &= m(x) - E[\hat{m}(x)] \\ &= m(x) - E\left[\sum_{i=1}^{n} Y_i\, w(x_i, x, h)\right] \\ &= m(x) - E\left[\sum_{i=1}^{n} (m(x_i) + \epsilon_i)\, w(x_i, x, h)\right] \\ &= m(x) - \sum_{i=1}^{n} m(x_i)\, w(x_i, x, h) \end{aligned}$$

Bootstrap

1. Estimate $\hat{P}$ for P_X (dist of data)
2. Generate simulation $\tilde{X}$ from $\hat{P}$, set $\tilde{T} = T(\tilde{X})$
3. Use simulated dist of $\tilde{T}$ to approximate P_T (dist of some statistic of X)

Model Based:
- best guess at P_X is $P_{X\hat{\theta}}$.
- get data and estimate $\hat{\theta}$ from X, then generate simulated data

Resampling Based:
- Sample with replacement from original data (size of data) n times, and then repeat this

Bootstrapping Regressions

- Hold x_i fixed, set $\tilde{y}_i = \hat{\mu}(x_i) + \hat{\epsilon}_i$, estimating noise dist
- Hold x_i fixed, set $\tilde{y}_i = \hat{\mu}(x_i) + \hat{\epsilon}_i$, sample from residuals
- resample (x_i, y_i) pairs

- simulate under null distribution

Significance Testing

- In R when creating a regression model, the significance tests of the βs are the following for each β_i
  H_{0i}: Assuming the model is right, $\beta_i = 0$
  H_{ai}: Assuming given model is right, $\beta_i \neq 0$
- uses Wald Test, with test statistic $T_j = \frac{\hat{\beta}_j - \beta_{null}}{\hat{se}(\hat{\beta}_j)}$
- So the significant βs will be:
  1. Have large true coefficients β_j
  2. Have large $V(X_j)$
  3. Have little correlation with other variables
  4. Also as n grows $\to \infty$ every coeff that isn't 0 will be signif

Confidence Set:

- either the true parameter is in set, model is wrong, or something improbable happened.
- For testing: reject $\theta = \theta_0$ if $\theta_0 \notin C$, retain if $\theta_0 \in C$